

FORUM

New Protocols Needed for Sharing Hydrometeorologic Data

The collection of hydrometeorological data is becoming privatized, and competition within the industrial sector (utilities, weather forecasting, etc.) is intensifying. Disruption in collection and distribution of these data will adversely affect both the science community and general public. The problem extends beyond California and the United States: governments worldwide are outsourcing data collection activities, and contractors recognize that restricting free access to hydrometeorological data increases profit margins.

The issue that must be raised, however, is whether restricting free public and governmental access to these data reduces water resources management efficiency as well as the accuracy and timeliness of flood forecasting, and delays the delivery of information to the public. Therefore, we propose a strategy that provides select users with real-time information for the benefit of the public, while protecting competitive interests within the private sector.

Electric utilities that manage hydrogeneration facilities have long been major collectors of hydrometeorologic data. In both the United States and other countries, the electric industry is being restructured to promote competition and decrease energy costs. In California, for example, deregulation will occur on April 1, 1998. Energy availability and cost are strongly influenced by current and future weather, snowpack, reservoir storage versus capacity, and river flow. Neighboring utilities who formerly shared these data will be at a competitive disadvantage unless they withhold data pertaining to their own hydrogeneration area.

Conversely, each power producer might gain a competitive edge by having ample information about their competitors' weather- and water-related operations. Similarly, research products sponsored by an industrial entity, that increase the accuracy of weather forecasts or runoff prediction, are unlikely to be shared willingly in California once April 1, 1998, has passed. A solution to the collaboration-versus-competition issue should be put in place, at least in California, by April 1998, to avoid disrupting a variety of state and federal programs. If such a solution becomes effective in California, perhaps it will be used as a model in other states and countries.

Open Data Repositories Versus Tiered Access

The Internet has greatly changed the way government agencies release their data to the public. Many federal and state agencies have invested heavily in Web technology to make both real-time and historic data available on their Web sites. For example, California's Department of Water Resources has established the California Data Exchange Center (CDEC) to act as an open repository for many types of data from many of California's agencies.

Open access, even within the federal system, is dependent on agency purpose. For example, the Columbia River daily flow forecasts made by the Bonneville Power Administration are available only to authorized internal users. Utilities and industrial concerns view the Web as an opportunity to market their products or services, but as a risk in terms of information control. Open repositories and data sharing will continue to be compromised as deregulation progresses.

A multi-tiered access control and a timed-release procedure, however, could solve many of the problems that are confronting data users. Repository data could be sequestered via restricted-access directories or via storage on entirely separate computers. By incorporating a time-delayed information release protocol, private or competing users would not risk having competitors gain access to their data until after the data were no longer relevant to current operations. A subset of special users such as a federal flood forecast center and a state seasonal runoff forecasting group, however, could still have immediate access to all data under emergency conditions.

Many groups currently have data access controls and public/private domain levels. The unique feature here is the interface between public and private data collection groups, the time requirements for access, and the disturbing threat that hydrometeorological data might be withheld from public safety user sectors (e.g., flood forecasting) unless competitive concerns are satisfied.

Expanding or modifying an existing centralized data center such as CDEC would be easier than instituting a new center because so many of the procedural and legal structures would be in place. Cooperating agencies enter the data or send forms for entry into the CDEC databases, and currently all of the data are available to all users. In the modified system, access to special, sensitive directories (or separate computer systems) would be protected by a password system. Public agency staff (National Weather Service, CDEC) or academic research staff would have access to most or all of the directories. Individual utilities, on the other hand, would have access to the public database and to the subdirectories containing

their own sensitive data, but not other utility's subdirectories. On a set schedule, the computer system manager (or an automated program) could move aged data out of the special directories and into the public data area.

Legal Issues

Computer capacity and legal restrictions associated with restricting access to public data need to be solved before groups like CDEC can have a role other than as an open repository. Federal agencies, for example, can institute a confidentiality agreement with an industrial partner who supplies special materials, formulas, or equipment. This agreement allows denial of Freedom of Information Act requests that pertain to a partner's proprietary information.

A similar mechanism needs to be made available for state agencies such as the California Department of Water Resources. Alternatively, the hydrometeorologic data might reside on a computer at an academic center. Also, waivers are needed to free each data contributor from liability due to use of data during natural disasters. Agreements establishing the special user category and censures for security violations should be established. An issue for special users is the mandate to present the public-safety products without explicitly disclosing the underlying, proprietary data.

Conclusions

Due to utility deregulation and government agency privatization, there is a need to implement protocols for sharing hydrometeorologic data. A centralized data system with multi-tiered access control and a timed-release procedure could solve many of the problems that are confronting data users.

As with any multigroup collaboration, the variety of missions, organizations, record keeping, and legal constraints will make the operation of the centralized data facility a challenge. The facility must formalize the mechanism for receiving data from cooperators and incorporating the data. Based on agreements with each cooperator, the data must be stored in a directory with the appropriate access control. Agreements should specify when the various types of data can be moved from restricted-access areas to public-access directories.

Many data collection and archiving systems exist, but they lack control protocols that allow special handling of proprietary data. We suggest that a dialogue begin immediately to establish guidelines for the establishment of such protocols.—*B. J. McGurk, Pacific Gas and Electric Company, Generation Portfolio Management, San Francisco, Calif., USA; and N. L. Miller, University of California, Lawrence Berkeley National Laboratory, Earth Sciences, Berkeley, Calif., USA*